

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平8-234796

(43)公開日 平成8年(1996)9月13日

(51)Int.Cl.<sup>6</sup>

G 1 0 L 9/18  
9/14

識別記号

府内整理番号

F I

G 1 0 L 9/18  
9/14

技術表示箇所  
E  
G  
J

審査請求 未請求 請求項の数 3 O L (全 8 頁)

(21)出願番号

特願平7-36467

(22)出願日

平成7年(1995)2月24日

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 金子 好之

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(72)発明者 中川 哲也

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(72)発明者 関根 英敏

東京都国分寺市東恋ヶ窪1丁目280番地

株式会社日立製作所中央研究所内

(74)代理人 弁理士 小川 勝男

最終頁に続く

(54)【発明の名称】 符号化音声の復号化器装置

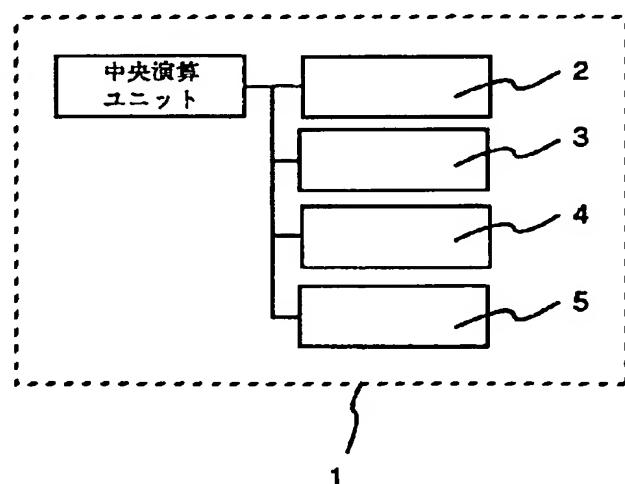
(57)【要約】

【目的】 言語学習機などパッケージ型の音声データを扱う機器に高能率符号化音声を適用するにあたり、それに好適なデータの格納方法およびその記憶装置ならびに復号化器装置を提供することにある。

【構成】 適応コードブックを用いる高能率符号化方式を用いて音声データを符号化し、その符号化データを記憶装置に格納する場合に、符号化時に用いた適応コードブックも同じ記憶装置に格納させる。

【効果】 復号化器装置は圧縮音声データと適応コードブックデータを記憶装置から取り出して復号化する。従って復号化器装置は、不特定の話者による音声データであっても良好な音質の音声を再生することができる。

図1



(2)

2

## 【特許請求の範囲】

【請求項1】情報と符号化音声データと話者の音声の特徴コードとが記憶された符号化器音声データの記憶装置と、上記記憶装置に記憶されている上記特徴コードに基づいて上記符号化音声データを復号化する復号化器装置を有することを特徴とする符号化音声の復号化器装置。

【請求項2】請求項1に記載の符号化音声の記憶装置およびこれを用いた復号化器装置において、前記特徴コードは前記符号化音声データに含まれる複数の話者の夫々の特徴に対応しており、前記記憶装置に複数個記憶されていることを特徴とする符号化音声の復号化器装置。

【請求項3】請求項1ないし2に記載の符号化音声の復号化器装置であって、前記特徴コードが、入力音声信号を予め定めた時間長のフレームに分割し、前記音声信号のスペクトル包絡を示すスペクトルパラメータを求めて出力し、前記フレームを予め定められた時間長のサブフレームに分割し、過去の音源から前記音声信号との誤差が最小になるよう長期予測パラメータを求めて出力し、前記サブフレームごとに駆動音源として予め用意されたコードブックの中から最適なコードベクトルを選択するCELP音声符号化方法における適応コードブックに記載されることを特徴とする符号化音声の復号化器装置。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、符号化した音声データの記憶装置および復号化器装置に係わり、特に複数の話者の声に対応させる場合に好適な符号化音声の記憶装置および復号化器装置に関する。

## 【0002】

【従来の技術】近年のコンピュータ技術の進展により、装置の小型化が進み、かつた処理能力が飛躍的に向上した。それに伴い、我々のまわりの情報をマルチメディア処理することが可能になって来ている。例えば、従来のテキスト、グラフィクスが支配的なコンピュータデータに音声や映像を援用したデータが容易に作成できる環境が整備されつつある。

【0003】上記のようなマルチメディア化に伴って、データの規模が飛躍的に増大しつつある。これは、音声や映像データが時系列的なデータ構造を有するための必然的な帰結である。しかしながら、これらのアナログデータについては、それらをデジタル変換した形態でそのまま格納するのではなく、冗長な信号を削除したりするなどの圧縮処理・符号化処理を施すのが通常である。これにより必要な記憶容量は数十分の一にまで低減することができるという効果が認められている。また、圧縮したデータを伝送するので伝送時の周波数の有効活用が可能になる。特に音声データについては、移動無線通信のデジタル化に向けて、上記周波数の有効活用の観点から、4 k b p s 以下の伝送レートの高能率音声符号化

方式が開発されている。

## 【0004】

【発明が解決しようとする課題】ところで、データの符号化・復号化に関する一般的な課題としては、原信号をいかに歪ませないかが最も重要なものである。あらゆる可能性のある信号に対してこの歪みを小さくするのは、かならずしも容易ではないが、特徴のある信号についてその歪みを小さくするように最適化するのは現実的な問題である。

【0005】例えば、特定話者の声に限って高能率符号化を適用するときには、その特定話者の声の特徴を十分考慮に入れた符号化が可能であればより高音質な復号化音を実現できる。実際、上記のような高能率音声符号化方式では、話者の特徴を考慮に入れるしきみとして、適応コードブックが採用されている。

【0006】従来は、上記の高能率符号化方式は移動無線などリアルタイムデータ用に開発されてきたのであって、上記のような適応性が十分活かされてきたわけではなかった。

【0007】パッケージ型の音声データについては話者が比較的特定できるので適応性を活用することが可能になる。

【0008】パッケージ型音声データを扱う機器の側からすると、それぞれのパッケージデータで話者が特定されても、多くの種類のパッケージデータに対応するには、話者は不特定とみなさざる得ない。一例としては、言語学習機が挙げられる。この場合、複数の言語に対応するといった機器の性格上、話者は不特定になる。

【0009】本発明の目的は、高能率符号化されたパッケージ型音声データを扱う機器において、話者が誰であってもそのデータの原音声の音質を最大限に引き出す構成を提供することにある。

【0010】本発明の目的は、上記のようにパッケージ型の音声データを扱う機器向けに高能率符号化音声を適用するにあたり、不特定多数の話者の音声であっても良好な音質の復号化音声を得るのに好適な構成でデータを格納する記憶装置を提供し、上記の復号化器装置が多くのパッケージ系データに対応可能とすることにある。

## 【0011】

【課題を解決するための手段】上記目的を達成するため、本発明では、話者の音声の符号化において話者の特徴となる部分を音声データと同時に記憶させる構成とした。具体的には、まず適応コードブックを用いる高能率符号化方式を用いて音声データを符号化する。次に、その符号化データを記憶装置に格納する場合に、符号化時に用いた適応コードブックも同じ記憶装置に格納するという構成をとる。さらに、データ復号化時には、復号化器装置は上記記憶装置から適応コードブック部分を読み出して、それをもとにデータを復号化するという構成とした。

(3)

3

## 【0012】

【作用】上記適応コードブックを用いる高能率符号化方式によれば、特定話者に限らずに常に良好な音質の圧縮音声データを得ることができる。さらにその適応コードブックデータは圧縮音声データとともに記憶装置に格納されており、復号化器は圧縮音声データと適応コードブックデータを上記記憶装置から取り出して復号化する。従って復号化器は、不特定の話者による音声データであっても前記の良好な音質の音声を再生することができる。

## 【0013】

【実施例】以下図面を用いて本発明の一実施例を説明する。図1に本発明による符号化音声の記憶装置を用いた復号化器の装置ブロック図を示す。

【0014】実施例における復号化器装置1は、中央演算ユニット(CPU)、データ表示部2、データ入力部3、データ主記憶部4、データ外部記憶装置5からなる。データ表示部2には対角約5インチの液晶ディスプレイを用いる。データ入力部3には、上記液晶ディスプレイに接着させた感圧式タッチパネルおよび簡単な押しボタン式スイッチを用いる。データ主記憶部4にはROMおよびRAMを、データ外部記憶装置5としては、カード型メモリを用いる。

【0015】図2は、4MBのメモリからなるデータ外部記憶装置5に格納されたデータの内容の一例である。外部記憶装置5には、少なくとも符号化された音声データ6と適応コードブックデータ7とが格納されている。符号化された音声データ6は、後述の高能率符号化により作成されたものであり、伝送レートは4kbps程度である。ここでは、約120分の音声データ6用に3.6MB、適応コードブックデータ7用に数10kBを充てた。

【0016】図3は、符号化器のブロック図である。本符号化器は、符号駆動線形予測(CELP)音声符号化方式に基づいて構成されている。音声の原データとしては所定の標本化周波数(通常8kHz)でA/D変換された音声信号101が入力される。音源の周期性を表す成分として適応コードブック108の出力である長期予測ベクトル110と、周期性以外の成分(ランダム性、雑音性)として利得112、113を乗じて加算した加重和114を駆動音源としている。

【0017】最適な駆動音源を得るためにコードブックの検索は以下のようにしてなされる。一般に、駆動音源を合成フィルタに入力して得られる合成音声が、原音声(入力音声)に一致するような駆動音源が得られればよいが、実際にはなんらかの誤差(量子化歪)を伴う。したがって、この誤差を最小化するように駆動音源を決定すればよい。その際、人間の聴覚特性との対応が良くなるように重み付けした誤差を用いるのが一般的である。

【0018】この聴覚重み付け誤差を評価するために、

4

駆動音源114は重み付け合成フィルタ105に入力され、重み付け合成音声116を得る。入力音声101も聴覚重み付けフィルタ104を通して重み付け入力音声115を得、重み付け合成音声116との差を取って重み付け誤差波形117を得る。なお、聴覚重み付けフィルタ104と重み付け合成フィルタ105のフィルタ係数は、あらかじめ入力音声101をLPC(線形予測)分析部102に入力して得られたLPCパラメータ103によって決められる。

【0019】重み付け誤差波形117は、2乗誤差計算部118において誤差評価区間にわたって2乗和を計算され、重み付け2乗誤差119が得られる。前述のように駆動音源は長期予測ベクトルと統計コードベクトルとの加重和であるから、駆動音源の決定は、各コードブックからどのコードベクトルを選択するかを決めるコードベクトル指標の決定に帰着する。すなわち、長期予測ラグ106とコードベクトル指標107を順次変えて重み付け2乗誤差119を算出し、誤差最小化部120において重み付け誤差が最小となるものを選択すればよい。

20 このような駆動音源決定法を「合成による分析法」と呼んでいる。

【0020】このようにして最適な駆動音源が決定したら、長期予測ラグ106、コードブック指標107、利得112、113、およびLPCパラメータ103をパラメータとして多重化部121において多重化したデータ122を外部記憶装置5に格納する。また、この時の駆動音源114を用いて適応コードブック108の状態が更新される。同一話者の音声を用いて、上記複数回の処理を繰り返すことにより、コードブックのトレーニングが完了する。もし外部記憶装置に格納する音声が複数人の音声であるならば、それら複数人の音声を使用したトレーニングが必要であることはいうまでもない。もちろん、上記のトレーニングが完了した後の適応コードブックを使用して音声データを符号化してもよい。

【0021】ここで適応コードブックの最終の状態のデータについても外部記憶装置5に記憶させる。これにより、以下に述べるような復号化器で、良好な音質の復号化音声を得ることが可能になる。なぜなら、話者固有の特徴を考慮にいれた最適な駆動音源が常に使用されるからである。

【0022】復号化器における処理は、図4に示された通りである。まず外部記憶装置から読み出された符号化データ222が、多重分離部221で各種パラメータに分離される。長期予測ラグ206に基づき適応コードブック208を検索し、長期予測ベクトル210を出力する。また、コードブック指標207に基づき統計コードブック209を検索し、音源ベクトル211を出力する。長期予測ベクトル210と音源ベクトル211にそれぞれの利得212と213を乗じ、加算した信号を駆動音源214として合成フィルタ230に入力する。合

(4)

5

成フィルタのフィルタ係数は、LPCパラメータ203によって決まる。ポストフィルタは必須ではないが、合成音声の主観的品質を改善するために多用され、その出力が出力音声232となる。ここでは、適応コードブックのデータは、データ外部記憶装置5に格納されているものを復号化装置の主記憶4に呼び出されて使用される。

【0023】なお、本発明においては、適応コードブックを用いる符号化復号化方式が前提となっているが、上記のような構成に限らないことはいうまでもない。例えば図5のような符号化器と図6のような復号化器の構成であってもよい。

【0024】この構成は、上記の例とは以下の点が異なる。図5に見られるように、音源として適応コードブックと雑音情報コードブックの他にパルス情報コードブックが加えられている。音響分類部からの入力情報すなわち入力音声の音響的特徴に基づいて、雑音源とパルス音源から適当なコードブックが検索処理の対象として選択可能としている。本発明では、話者の声の特徴がその符号化データとともに記憶装置に記録されていることが本質であって、符号化・復号化の詳細なアルゴリズムの差異に影響されるものではない。

【0025】次に、本発明による符号化音声の記憶装置およびこれを用いた復号化器装置を言語学習機に応用した例を示す。図7の10～15は6か国の言語の符号化音声データとそれぞれの話者A～Fの適応コードブックデータを格納した外部記憶装置である。復号化器装置は、上記外部記憶装置10～15いずれにも対応可能であり、それぞれのコードブックデータを呼び出して復号化するのでそれらの話者の特徴に適応した音声復号ができる。

【0026】上記では、複数の話者が含まれる音声データの作成には、それぞれの話者のデータを用いて单一のコードブックをトレーニングする例について述べた。図8に、複数の話者のそれぞれに適応コードブックデータを備える例を示す。外部記憶装置16の内部には3種類の適応コードブックG、H、Iが備わり、それぞれが異なる話者データGd、Hd、Idに対応している。適応コードブックG、H、Iは、それぞれ話者Gd、Hd、

6

Idの音声によりトレーニングが済ませてある。

【0027】復号化器装置側でデータを読み込み復号化を行うときに、それぞれの組合せを使用するようにすれば、最適な復号化音声を得ることができる。

【0028】なお本発明は上記実施例によって限定されるものではない。第一に、復号化器装置の構成は上記に限ったものではない。例えば、入力装置としては、タッチパネルと押しボタン式スイッチの組み合わせに限らない。さらには、外部記憶装置としては、カード型メモリに限らずCD-ROMのような光学読み出し式メモリディスクであってもよい。第二に、用途についても言語学習機に限らず例えば異種言語間の通訳器でもよく、音楽再生装置でもよい。

【0029】

【発明の効果】以上本発明によれば次のような効果がある。高能率符号化方式によって作成した音声データの復号化時に、その音声データに付随して記憶された話者の声の特徴に関するコードブックデータも記憶装置から取り出されて援用されるので特定話者に限らずに常に良好な音質の音声データを得ることができる。その結果、同一の復号化器装置を用いて、圧縮符号化音声データの活用を図ることができる。

【図面の簡単な説明】

【図1】本発明による符号化音声の記憶装置を用いた復号化器装置のブロック図。

【図2】本発明による符号化音声の外部記憶装置の構成を説明するための図。

【図3】符号化器の動作ブロック図。

【図4】復号化器の動作ブロック図。

【図5】別の符号化器の動作ブロック図。

【図6】別の復号化器の動作ブロック図。

【図7】本発明による言語学習機を説明するための図。

【図8】別の構成による外部記憶装置を説明するための図。

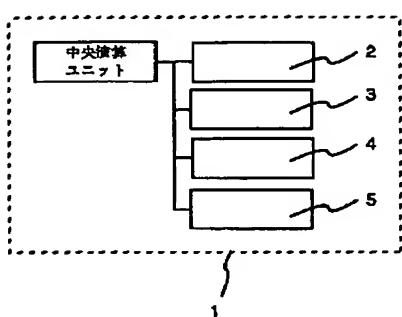
【符号の説明】

1…復号化器装置、2…データ表示部、3…データ入力部、4…データ主記憶部、5…データ外部記憶装置、6…符号化音声データ、7…適応コードブックデータ。

(5)

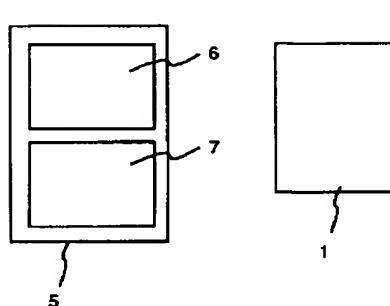
【図1】

図1



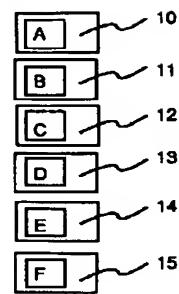
【図2】

図2

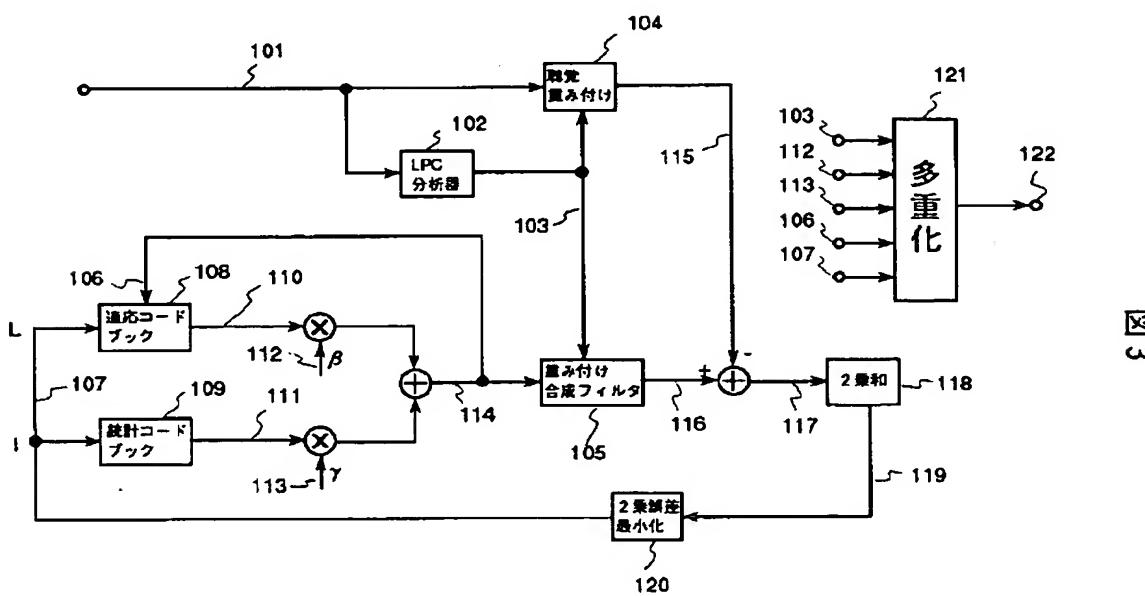


【図7】

図7

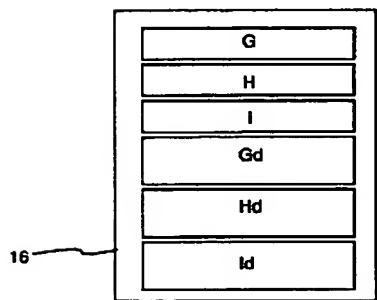


【図3】



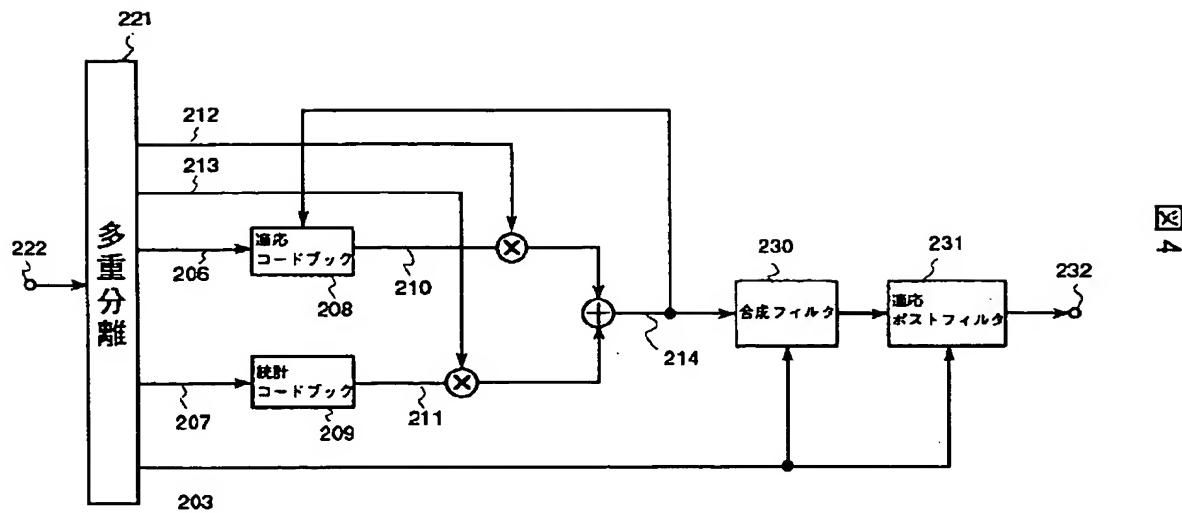
【図8】

図8



(6)

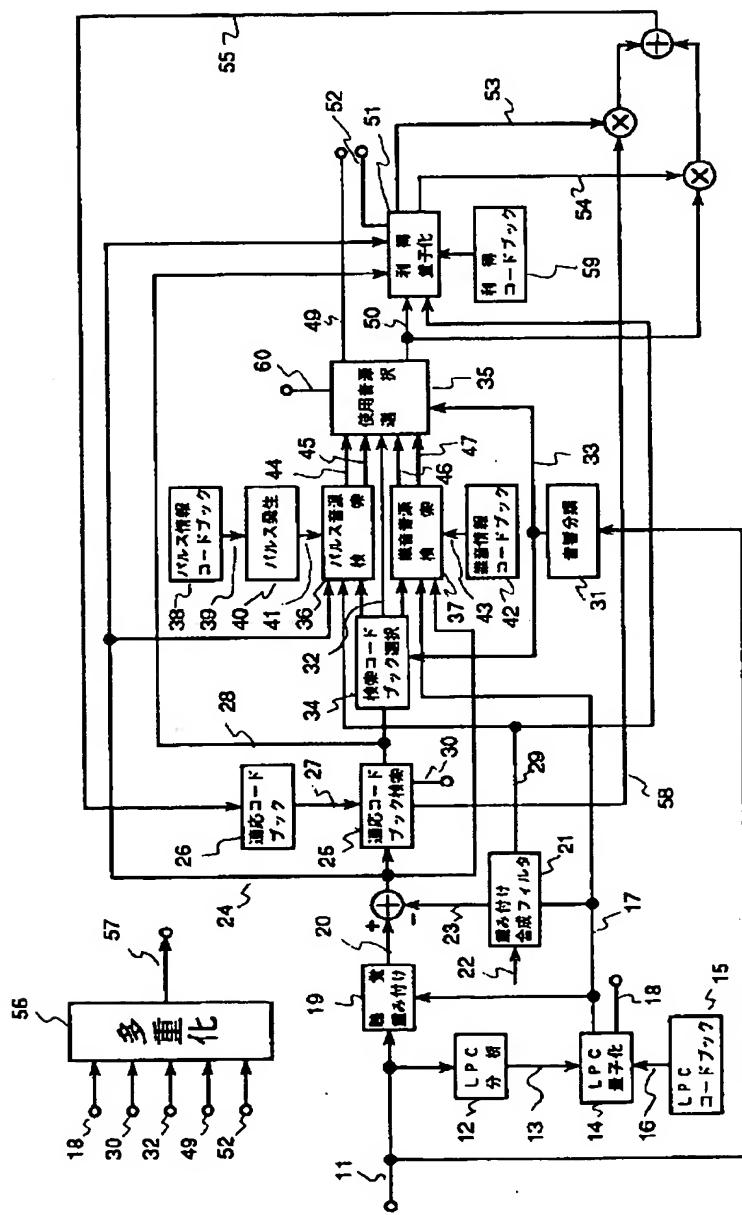
[図4]



(7)

【図5】

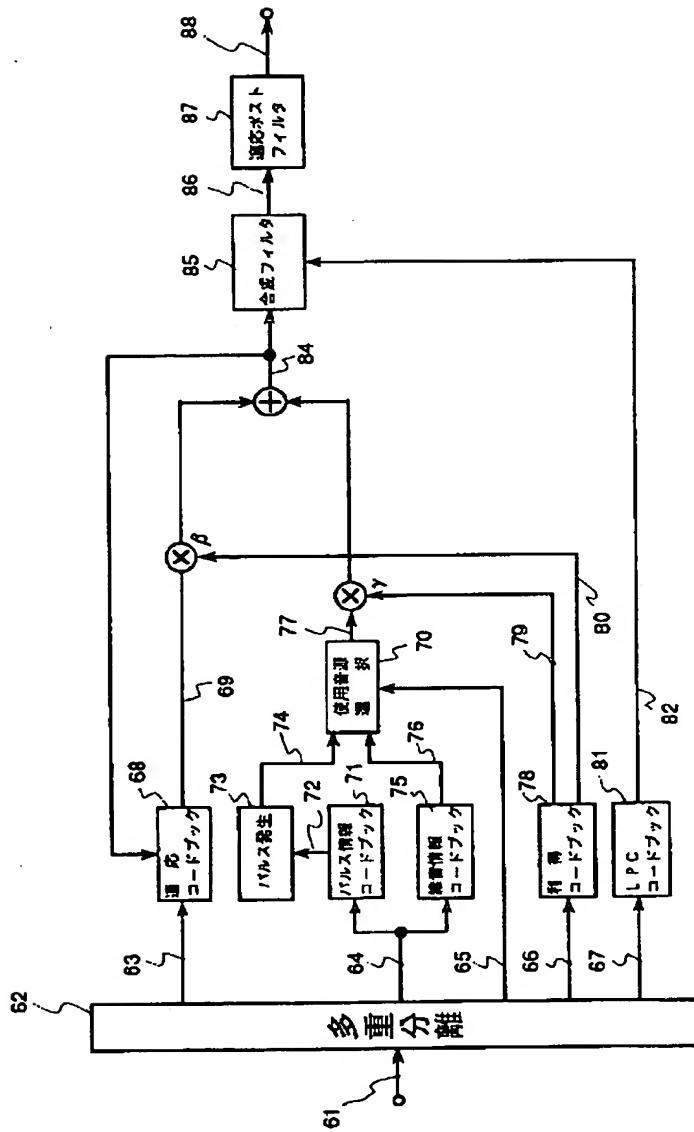
図5



(8)

【图6】

6



## フロントページの続き

(72) 発明者 荻路 憲治

神奈川県横浜市戸塚区吉田町292番地 株  
式会社日立製作所マルチメディアシステム  
事業部内

## PATENT ABSTRACTS OF JAPAN

(11) Publication number : 08-234796

(43) Date of publication of application : 13.09.1996

(51)Int.CI. G10L 9/18  
G10L 9/14

(21) Application number : 07-036467

(71)Applicant : HITACHI LTD

(22) Date of filing : 24.02.1995

(72)Inventor : KANEKO YOSHIYUKI  
NAKAGAWA TETSUYA  
SEKINE HIDETOSHI  
OGIJI KENJI

(54) DECODER DEVICE FOR ENCODED VOICE

(57) Abstract:

**PURPOSE:** To always obtain voice data having a satisfactory sound quality without being limited to a specific speaker and to make use of compression encoded voice data by using the same decoder device by making a constitution making parts becoming the feature of a speaker to be stored at the same time of storing voice data as to the encoding of the voice of the speaker.

**CONSTITUTION:** A decoder device 1 is constituted of a central processing unit (a CPU), a data display part 2, a data input part 3, a data main storage part 4 and a data external storage part 5. In the data external storage part 5, at least encoded voice data 6 and adaptive code book data 7 are stored. Then, voice data are encoded by a high efficiency encoding system while using the adaptive code book data 7. Next, in the case of storing encoded data in the storage part 5, the adaptive code book used at the time of the encoding are also stored in the same storage device 5. Next, at the time of the data decodings, the decoder device 1 reads out the adaptive cod book part from the device 5 and then decodes data based on the readout part.

## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

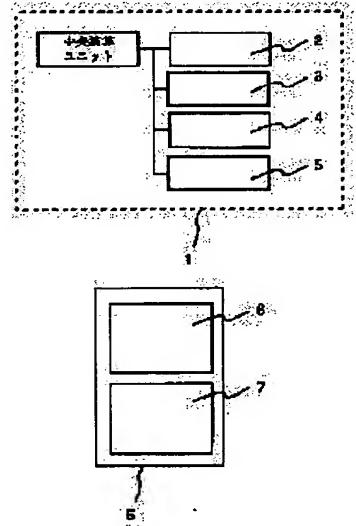
[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]



**\* NOTICES \***

**JPO and NCIPI are not responsible for any damages caused by the use of this translation.**

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. \*\*\*\* shows the word which can not be translated.

3. In the drawings, any words are not translated.

---

**CLAIMS**

---

**[Claim(s)]**

**[Claim 1]** Decryption machine equipment of the coding voice characterized by having the store of the encoder voice data with which information, coding voice data, and the feature code of a speaker's voice were memorized, and decryption machine equipment which decrypts the above-mentioned coding voice data based on the above-mentioned feature code memorized by the above-mentioned store.

**[Claim 2]** It is decryption machine equipment of the coding voice characterized by supporting each description of two or more speakers by whom said feature code is contained in said coding voice data in the store of coding voice according to claim 1, and the decryption machine equipment using this, and more than one being memorized by said store.

**[Claim 3]** It is decryption machine equipment of claim 1 thru/or coding voice given in 2. Said feature code Divide into the frame of the time amount length which defined the input sound signal beforehand, and it outputs in quest of the spectrum parameter which shows the spectral envelope of said sound signal. Divide into the subframe of the time amount length which was able to define said frame beforehand, and in quest of a long-term prediction parameter, it outputs so that an error with said sound signal may become min from the past sound source. Decryption machine equipment of the coding voice characterized by what is indicated by the adaptation code book in the CELP voice coding approach which chooses the optimal code vector out of the code book beforehand prepared as a drive sound source for said every subframe.

---

**[Translation done.]**

**\* NOTICES \***

**JPO and NCIPI are not responsible for any damages caused by the use of this translation.**

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. \*\*\*\* shows the word which can not be translated.

3. In the drawings, any words are not translated.

---

**DETAILED DESCRIPTION**

---

**[Detailed Description of the Invention]**

**[0001]**

**[Industrial Application]** With respect to the storage and decryption machine equipment of voice data which were encoded, when making this invention correspond to two or more speakers' voice especially, it relates to the storage and decryption machine equipment of suitable coding voice.

**[0002]**

**[Description of the Prior Art]** By progress of computer technology in recent years, the miniaturization of

equipment progressed and the throughput improved by leaps and bounds again. In connection with it, it is becoming possible to carry out multimedia processing of our information on surrounding. For example, the environment which the conventional text and the data with which graphics used voice and an image for dominant computer data can create easily is being improved.

[0003] The scale of data is increasing by leaps and bounds with the above multimedia-izing. This is an inevitable conclusion for having DS with serial voice and image data. However, usually compression processing / coding processing of deleting a redundant signal about these analog data, rather than storing as it is with the gestalt which carried out digital conversion of them is performed. Since it becomes possible to reduce required storage capacity to 1/dozens by this, the effectiveness that an equipment scale and cost can be reduced is accepted. Moreover, since compressed data are transmitted, effective use of the frequency at the time of transmission is attained. The high efficiency voice coding method of the transmission rate of 4 or less kbps is especially developed just by voice data from a viewpoint of effective use of the above-mentioned frequency towards digitization of \*\* and a mobile radio communication link.

[0004]

[Problem(s) to be Solved by the Invention] By the way, as a general technical problem about coding and a decryption of data, it is most important how the HARASHIN number is not made distorted. Although it is not necessarily easy to make this distortion small to a signal with all possibility, it is a realistic problem to optimize so that that distortion may be made small about a characteristic signal.

[0005] if coding which took the description of the voice of the specified speaker into consideration enough is possible when [ of a specified speaker ] applying high efficiency coding only within voice -- more -- high -- a tone quality decryption sound is realizable. By the above high efficiency voice coding methods, the adaptation code book is actually adopted as structure which takes a speaker's description into consideration.

[0006] Conventionally, the above-mentioned low bit rate coding method had not been developed for [ , such as a mobile radio, ] real-time data, and the above adaptability had not necessarily been harnessed enough.

[0007] Since a speaker can specify comparatively about the voice data of a package mold, it becomes possible to utilize adaptability.

[0008] Even if the speaker is specified by each package data, considering the device side treating package mold voice data, in order to correspond to many kinds of package data, a speaker cannot regard it as unspecified. A linguistics learning machine is mentioned as an example. In this case, a speaker becomes unspecified on the character of the device of corresponding to two or more language.

[0009] In the device treating the package mold voice data by which high efficiency coding was carried out, the purpose of this invention is to offer the configuration which pulls out the tone quality of the Hara voice of the data to the maximum extent, even if a speaker is whom.

[0010] In applying high-efficiency-coding voice for the devices which treat the voice data of a package mold as mentioned above, even if the purpose of this invention is many and unspecified speakers' voice, it offers the store which stores data with a suitable configuration to obtain the decryption voice of good tone quality, and it is for above decryption machine equipment to enable correspondence to many package system data.

[0011]

[Means for Solving the Problem] In order to attain the above-mentioned purpose, in this invention, it considered as the configuration which makes voice data and coincidence memorize the part which serves as a speaker's description in coding of a speaker's voice. Specifically, voice data is first encoded using the low bit rate coding method using an adaptation code book. Next, when it stores the coded data in a store, the configuration of also storing in the same store the adaptation code book used at the time of coding is taken. Furthermore, at the time of a data decryption, decryption machine equipment read the adaptation code book part from the above-mentioned store, and considered it as the configuration of decrypting data based on it at it.

[0012]

[Function] According to the low bit rate coding method using the above-mentioned adaptation code book, the compression voice data of not only a specified speaker but always good tone quality can be obtained. Furthermore, the adaptation code book data is stored in the store with compression voice data, compression voice data and adaptation code book data are picked out from the above-mentioned store, and a decryption machine decrypts them. Therefore, even if a decryption machine is the voice data by the unspecified speaker, it can reproduce the voice of the aforementioned good tone quality.

[0013]

[Example] One example of this invention is explained using a drawing below. The equipment block diagram of the decryption machine using the store of the coding voice by this invention is shown in drawing 1.

[0014] The decryption machine equipment 1 in an example consists of a central arithmetic unit (CPU), the data display section 2, the data input section 3, a data primary storage 4, and data external storage 5. The liquid crystal display of about 5 inches of vertical angles is used for the data display section 2. The pressure-sensitive type touch panel and the easy push button type switch which were pasted up on the above-mentioned liquid crystal display are used for the data input section 3. As data external storage 5, card mold memory is used for the data primary storage 4 for ROM and RAM.

[0015] Drawing 2 is an example of the contents of the data stored in the data external storage 5 which consists of 4MB of memory. The voice data 6 and the adaptation code book data 7 which were encoded at least are stored in external storage 5. The encoded voice data 6 is created by the below-mentioned high efficiency coding, and a transmission rate is 4kbps extent. Here, several 10 kB(s) were allotted to the voice data 6 for about 120 minutes 3.6MB and for the adaptation code book data 7.

[0016] Drawing 3 is the block diagram of an encoder. This encoder was constituted based on the sign drive linear prediction (CELP) voice coding method, and is. The sound signal 101 by which A/D conversion was carried out with the predetermined sampling frequency (usually 8kHz) as audio original data is inputted. The weighted sum 114 which multiplied by them and added gains 112 and 113 as the long-term prediction vector 110 which is the output of the adaptation code book 108 as a component showing the periodicity of a sound source, and components other than periodicity (random nature and noise nature) is made into the drive sound source.

[0017] Retrieval of the code book for acquiring the optimal drive sound source is the following, and is made and made. Although a drive sound source whose synthesized speech which inputs a drive sound source into a synthetic filter, and is obtained generally corresponds with the Hara voice (input voice) should just be acquired, it is accompanied by a certain error (quantumization noise) in fact. Therefore, what is necessary is just to determine that a drive sound source will minimize this error. It is common in that case to use the error which carried out weighting so that correspondence with human being's acoustic-sense property might become good.

[0018] In order to evaluate this acoustic-sense weighting error, the drive sound source 114 is inputted into the weighting composition filter 105, and obtains the weighting synthesized speech 116. the input voice 101 -- the acoustic-sense weighting filter 104 -- letting it pass -- the weighting input voice 115 -- obtaining -- a difference with the weighting synthesized speech 116 -- taking -- a weighting error wave -- 117 is obtained. In addition, the filter factor of the acoustic-sense weighting filter 104 and the weighting composition filter 105 is decided with the LPC parameter 103 which inputted the input voice 101 into the LPC (linear prediction) analyzor 102 beforehand, and was obtained.

[0019] Weighting error wave 117 has the square sum calculated over the error evaluation section in the square error count section 118, and the weighting square error 119 is acquired. As mentioned above, since a drive sound source is the weighted sum of a long-term prediction vector and a statistics code vector, the decision of a drive sound source results in the decision of the code vector index which decides which code vector to choose from each code book. Namely, what is necessary is to compute the weighting square error 119 by changing the long-term prediction lug 106 and the code vector index 107 one by one, and just to choose that from which a weighting error serves as min in the error minimization section 120. Such a drive sound-source determining method is called "the analysis method by composition."

[0020] Thus, if the optimal drive sound source is determined, the long-term prediction lug 106, the code book index 107, gains 112 and 113, and the data 122 multiplexed in the multiplexing section 121 by making the LPC parameter 103 into a parameter are stored in external storage 5. Moreover, the condition of the adaptation code book 108 is updated using the drive sound source 114 at this time. Training of a code book is completed by repeating processing of the above-mentioned multiple times using the same speaker's voice. If the voice stored in external storage is two or more persons' voice, it cannot be overemphasized that training which used these two or more people's voice is required. Of course, voice data may be encoded using the adaptation code book after the above-mentioned training is completed.

[0021] External storage 5 is made to memorize also about the data of the last condition of an adaptation code book here. It enables this to obtain the decryption voice of good tone quality with a decryption vessel which is described below. It is because the optimal drive sound source which took the description of a speaker proper into consideration is always used.

[0022] The processing in a decryption machine is as having been shown in drawing 4. The coded data 222 first read from external storage is divided into various parameters in the demultiplexing section 221. The adaptation code book 208 is searched based on the long-term prediction lug 206, and the long-term prediction vector 210 is outputted. Moreover, the statistics code book 209 is searched based on the code book index 207, and the sound-source vector 211 is outputted. The long-term prediction vector 210 and the sound-source vector 211 are multiplied by each gain 212 and 213, and it inputs into the synthetic filter 230 by making the added signal into the drive sound source 214 at them. The filter factor of a synthetic filter is decided with the LPC parameter 203. Although a postfilter is not indispensable, it is used abundantly in order to improve the subjective quality of synthesized speech, and the output serves as the output voice 232. Here, the data of an adaptation code book are used in what is stored in the data external storage 5, being called to the primary storage 4 of decryption machine equipment.

[0023] In addition, in this invention, although the coding decryption method using an adaptation code book is the requisite, it cannot be overemphasized that it does not restrict to the above configurations. For example, you may be the configuration of an encoder like drawing 5, and a decryption machine like drawing 6.

[0024] As for this configuration, the following points differ from the above-mentioned example. The pulse information code book other than an adaptation code book and a noise information code book is added as a sound source so that drawing 5 may see. Based on the acoustical description of the input from the sound classification section, i.e., input voice, the suitable code book supposes that it is selectable as an object of retrieval processing from the noise source and the pulse sound source. In this invention, it is essence that the description of a speaker's voice is recorded on storage with the coded data, and it is not influenced by the difference in the detailed algorithm coding and a decryption.

[0025] Next, the example which applied the storage of the coding voice by this invention and the decryption machine equipment using this to the linguistics learning machine is shown. 10-15 of drawing 7 are the external storage which stored the coding voice data of the language of six nations, and the adaptation code book data of each speaker A-F. decryption machine equipment -- the above-mentioned external storage 10-15 -- it can respond to all, and since it decrypts by calling each code book data, the voice decode which was adapted for those speakers' description can be performed.

[0026] The above described the example which trains a single code book using each speaker's data to creation of the voice data with which two or more speakers are contained. The example which equips drawing 8 R>8 with adaptation code book data at each of two or more speakers is shown. The interior of external storage 16 is equipped with three kinds of adaptation code books G, H, and I, and the speaker data Gd, Hd, and Id with which each differs are supported. As for the adaptation code books G, H, and I, training is finished with Speakers' Gd, Hd, and Id voice, respectively.

[0027] If each combination is used when decrypting by reading data by the decryption machine equipment side, the optimal decryption voice can be obtained.

[0028] In addition, this invention is not limited by the above-mentioned example. It is not what restricted the configuration of decryption machine equipment in the first place above. For example, as an input

device, it does not restrict to the combination of a touch panel and a push button type switch. Furthermore, as external storage, you may be an optical read-out type memory disk not only like card mold memory but CD-ROM. The interpreter machine for example, not only between a linguistics learning machine but different-species language may be used also about the second application, and a music regenerative apparatus is sufficient.

[0029]

[Effect of the Invention] According to this invention, there is the following effectiveness above. Since the code book data about the description of a speaker's voice memorized along with the voice data are also picked out from a store and are applied at the time of a decryption of the voice data created with the low bit rate coding method, the voice data of not only a specified speaker but always good tone quality can be obtained. Consequently, the activity of compression coding voice data can be aimed at using the same decryption machine equipment.

---

[Translation done.]

\* NOTICES \*

JPO and NCIPI are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] The block diagram of the decryption machine equipment using the store of the coding voice by this invention.

[Drawing 2] Drawing for explaining the configuration of the external storage of the coding voice by this invention.

[Drawing 3] The block diagram of an encoder of operation.

[Drawing 4] The block diagram of a decryption machine of operation.

[Drawing 5] The block diagram of another encoder of operation.

[Drawing 6] The block diagram of another decryption machine of operation.

[Drawing 7] Drawing for explaining the linguistics learning machine by this invention.

[Drawing 8] Drawing for explaining the external storage by another configuration.

[Description of Notations]

1 [ -- A data primary storage, 5 / -- Data external storage, 6 / -- Coding voice data, 7 / -- Adaptation code book data. ] -- Decryption machine equipment, 2 -- The data display section, 3 -- The data input section, 4

---

[Translation done.]